

FIG. 1A

EcoRIa 24 39 54
GAATTCCGG TGC AGG ACG AAG CTG TTC TGG ATT TCT TAC AGT GAT GGG GAC CAG
 Arg Thr Lys Leu Phe Trp Ile Ser Tyr Ser Asp Gly Asp Gln

69 84 99
 TGT GCC TCA AGT CCA TGC CAG AAT GGG GGC TCC TGC AAG GAC CAG CTC CAG TCC
 Cys Ala Ser Ser Pro Cys Gln Asn Gly Gly Ser Cys Lys Asp Gln Leu Gln Ser

114 129 144 159
 TAT ATC TGC TTC TGC CTC CCT GCC TTC GAG GGC CGG AAC TGT GAG ACG CAC AAG
 Tyr Ile Cys Phe Cys Leu Pro Ala Phe Glu Gly Arg Asn Cys Glu Thr His Lys

174 189 204 Pst Ia
 GAT GAC CAG CTG ATC TGT GTG AAC GAG AAC GGC GGC TGT GAG CAG TAC TGC AGT
 Asp Asp Gln Leu Ile Cys Val Asn Glu Asn Gly Gly Cys Glu Gln Tyr Cys Ser

219 234 249 264
 GAC CAC ACG GGC ACC AAG CGC TCC TGT CGG TGC CAC GAG GGG TAC TCT CTG CTG
 Asp His Thr Gly Thr Lys Arg Ser Cys Arg Cys His Glu Gly Tyr Ser Leu Leu

279 294 309 324
 GCA GAC GGG GTG TCC TGC ACA CCC ACA GTT GAA TAT CCA TGT GGA AAA ATA CCT
 Ala Asp Gly Val Ser Cys Thr Pro Thr Val Glu Tyr Pro Cys Gly Lys Ile Pro

Xba I 339 354 369
 ATT CTA GAA AAA AGA AAT GCC AGC AAA CCC CAA GGC CGA ATT GTG GGG GGC AAG
 Ile Leu Glu Lys Arg Asn Ala Ser Lys Pro Gln Gly Arg Ile Val Gly Gly Lys

384 399 414 429
 GTG TGC CCC AAA GGG GAG TGT CCA TGG CAG GTC CTG TTG TTG GTG AAT GGA GCT
 Val Cys Pro Lys Gly Glu Cys Pro Trp Gln Val Leu Leu Leu Val Asn Gly Ala

444 459 474
 CAG TTG TGT GGG GGG ACC CTG ATC AAC ACC ATC TGG GTG GTC TCC GCG GCC CAC
 Gln Leu Cys Gly Gly Thr Leu Ile Asn Thr Ile Trp Val Val Ser Ala Ala His

489 504 519 534
 TGT TTC GAC AAA ATC AAG AAC TGG AGG AAC CTG ATC GCG GTG CTG GGC GAG CAC
 Cys Phe Asp Lys Ile Lys Asn Trp Arg Asn Leu Ile Ala Val Leu Gly Glu His

549 564 579 594
 GAC CTC AGC GAG CAC GAC GGG GAT GAG CAG AGC CGG CGG GTG GCG CAG GTC ATC
 Asp Leu Ser Glu His Asp Gly Asp Glu Gln Ser Arg Arg Val Ala Gln Val Ile

609 Sma I 624 639
 ATC CCC AGC ACG TAC GTC CCG GGC ACC ACC AAC CAC GAC ATC GCG CTG CTC CGC
 Ile Pro Ser Thr Tyr Val Pro Gly Thr Thr Asn His Asp Ile Ala Leu Leu Arg

654 669 684 699
 CTG CAC CAG CCC GTG GTC CTC ACT GAC CAT GTG GTG CCC CTC TGC CTG CCC GAA
 Leu His Gln Pro Val Val Leu Thr Asp His Val Val Pro Leu Cys Leu Pro Glu

| | | |
|--|-------------------------------------|-----------------------------|
| 714 | 729 | 744 |
| CGG ACG TTC TCT GAG AGG ACG CTG GCC TTC GTG CGC TTC TCA TTG GTC AGC GGC | | |
| Arg Thr Phe Ser Glu Arg Thr Leu Ala Phe Val Arg Phe Ser Leu Val Ser Gly | | |
| 759 | 774 | Nar I |
| TGG GGC CAG CTG CTG GAC CGT GGC GCC | 789 | 804 |
| Trp Gly Gln Leu Leu Asp Arg Gly Ala | Thr Ala | Leu Glu Leu Met Val Leu Asn |
| 819 | 834 | Pst Ib |
| GTG CCC CGG CTG ATG ACC CAG GAC TGC | 849 | 864 |
| Val Pro Arg Leu Met Thr Gln Asp Cys | Leu Gln Gln Ser Arg Lys Val Gly Asp | |
| 879 | 894 | 909 |
| TCC CCA AAT ATC ACG GAG TAC ATG TTC TGT GCC GGC TAC TCG GAT GGC AGC AAG | | |
| Ser Pro Asn Ile Thr Glu Tyr Met Phe Cys Ala Gly Tyr Ser Asp Gly Ser Lys | | |
| 924 | 939 | 954 |
| GAC TCC TGC AAG GGG GAC AGT GGA GGC CCA CAT GCC ACC CAC TAC CGG GGC ACG | | 969 |
| Asp Ser Cys Lys Gly Asp Ser Gly Gly Pro His Ala Thr His Tyr Arg Gly Thr | | |
| 984 | 999 | 1014 |
| TGG TAC CTG ACG GGC ATC GTC AGC TGG GGC CAG GGC TGC GCA ACC GTG GGC CAC | | |
| Trp Tyr Leu Thr Gly Ile Val Ser Trp Gly Gln Gly Cys Ala Thr Val Gly His | | |
| 1029 | 1044 | 1059 |
| TTT GGG GTG TAC ACC AGG GTC TCC CAG TAC ATC GAG TGG CTG CAA AAG CTC ATG | TaqI | 1074 |
| Phe Gly Val Tyr Thr Arg Val Ser Gln Tyr Ile Glu Trp Leu Gln Lys Leu Met | | |
| 1089 | 1104 | 1119 |
| CGC TCA GAG CCA CGC CCA GGA GTC CTC CTG CGA GCC CCA TTT CCC TAG CCCAGCAGCC | | 1138 |
| Arg Ser Glu Pro Arg Pro Gly Val Leu Leu Arg Ala Pro Phe Pro | | |
| 1148 | 1158 | 1168 |
| CTGGCCTGTG GAGAGAAAGC CAAGGCTGCG TCGAACTGTC | 1178 | 1188 |
| | 1198 | PstIc |
| | | 1208 |
| | | TTCTTCTGCA |
| 1218 | 1228 | 1238 |
| GTTAATGGGG TAGAGGAGGG CATGGGAGGG AGGGAGAGGT | 1248 | 1258 |
| | 1268 | 1278 |
| | | AAACAGAGAG |
| 1288 | 1298 | 1308 |
| AGACAGAGAC AGAGAGAGAC TGAGGGAGAG ACTCTGAGGA | 1318 | 1328 |
| | 1338 | 1348 |
| | | AGACTCAAAG AGACTCCAAG |
| 1358 | 1368 | 1378 |
| ATTCAAAGAG ACTAATAGAG ACACAGAGAT GGAATAGAAA | 1388 | 1398 |
| | 1408 | 1418 |
| | | CAGAGGCAGA CAGGCGCTGG |
| 1428 | 1438 | 1448 |
| ACAGAGGGGC AGGGGAGTGC CAAGGTTGTC CTGGAGGCAG | 1458 | 1468 |
| | 1478 | 1488 |
| | | TGAGCCTCCT TACCTCCCTT |

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|
| 1498 | 1508 | 1518 | 1528 | 1538 | 1548 | 1558 |
| CAGCCAAGCC | CCACCTGCAC | GTGATCTGCT | GGCCCTCAGG | CTGCTGCTCT | GCCTTCATTG | CTGGAGACAG |
| 1568 | 1578 | 1588 | 1598 | 1608 | 1618 | 1628 |
| TAGAGGCATG | ACACACATGG | ATGCACACAC | ACACACGCCA | TGCACACACA | CAGAGATATG | CACACACACG |
| 1638 | 1648 | 1658 | 1668 | 1678 | 1688 | 1698 |
| GATGCACACA | CAGATGGTCA | CACAGAGTAC | GCAAACACAC | CGATGCACAC | GCACATAGAG | ATATGCACAC |
| 1708 | 1718 | 1728 | 1738 | 1748 | 1758 | 1768 |
| ACAGATGCAC | ACACAGATAT | ACACATGGAG | TGCACGCACA | TGCCAATGCA | CGCACACATC | AGTGCACACG |
| 1778 | 1788 | 1798 | 1808 | 1818 | 1828 | 1838 |
| GATGCACAGA | GATATGCACA | CACCGATGTG | CGCACACACA | GATATGCACA | CACATGGATG | AGCACACACA |
| 1848 | 1858 | 1868 | 1878 | 1888 | 1898 | 1908 |
| CACCAAGTGC | GCACACACAC | CGATGTACAC | ACAGATGCAC | ACACAGATGC | ACACACACCG | ATGCTGACTC |
| 1918 | 1928 | 1938 | 1948 | 1958 | 1968 | 1978 |
| CATGTGTGCT | GTCCTCTGAA | GGCGGTTGTT | TAGCTCTCAC | TTTTCTGGTT | CTTATCCATT | ATCATCTTCA |
| 1988 | 1998 | 2008 | 2018 | 2028 | 2038 | 2048 |
| CTTCAGACAA | TTCAGAAGCA | TCACCATGCA | TGGTGGCGAA | TGCCCCCAAA | CTCTCCCCCA | AATGTATTTT |
| 2058 | 2068 | 2078 | 2088 | 2098 | 2108 | 2118 |
| TCCCTTCGCT | GGGTGCCGGG | CTGCACAGAC | TATTCCCCAC | CTGCTTCCCA | GCTTCACAAT | AAACGGCTGC |
| 2128 | 2138 | 2148 | 2158 | 2168 | EcoRIb | |
| GTCTCCTCGC | AAAAAAAAAA | AAAAAAAAAA | AAAAAAAAAA | AAAAAAAAAA | AAGGAATTC | |

FIG. 1B

-60
 TCAACAGGCAGGGGCAGCACTGCAGAGATTTTCATCATGGTCTCCCAGGCCCTCAGGCTCCTC
 10 20 30 40 50 60
 MetValSerGlnAlaLeuArgLeuLeu
 -50 -40 ↓
 CysLeuLeuLeuGlyLeuGlnGlyCysLeuAlaAlaGlyGlyValAlaLysAlaSerGlyGly
 TGCCTTCTGCTTGGGCTTCAGGGCTGCCTGGCTGCAGGCGGGTTCGCTAAGGCCTCAGGAGGA
 70 80 90 100 110 120
 -30 -20 ↓
 GluThrArgAspMetProTrpLysProGlyProHisArgValPheValThrGlnGluGlu
 GAAACACGGGACATGCCGTGGAAGCCGGGCTCACAGAGTCTTCGTAACCCAGGAGGAA
 130 140 150 160 170 180
 -10 -1 +1 +10
 AlaHisGlyValLeuHisArgArgArgAlaAsnAlaPheLeuGluGluLeuArgPro
 GCCACGGCGTCTGCACCGGCGCGCGGCCAACGCGTTCCTGGAGGAGCTGCGGCCG
 190 200 210 220 230 240
 +20 +30
 GlySerLeuGluArgGluCysLysGluGluGlnCysSerPheGluGluAlaArgGluIle
 GGCTCCCTGGAGAGGGAGTGCAAGGAGGAGCAGTGCTCCTTCGAGGAGGCCCGGAGATC
 250 260 270 280 290 300
 +40 +50
 PheLysAspAlaGluArgThrLysLeuPheTrpIleSerTyrSerAspGlyAspGlnCys
 TTCAAGGACGCGGAGAGGACGAAGCTGTTCTGGATTTCTTACAGTGATGGGGACCAAGTGT
 310 320 330 340 350 360
 +60 +70
 AlaSerSerProCysGlnAsnGlyGlySerCysLysAspGlnLeuGlnSerTyrIleCys
 GCCTCAAGTCCATGCCAGAATGGGGGCTCCTGCAAGGACCAGCTCCAGTCCATATATCTGC
 370 380 390 400 410 420
 +80 +90
 PheCysLeuProAlaPheGluGlyArgAsnCysGluThrHisLysAspAspGlnLeuIle
 TTCTGCCTCCCTGCCTTCGAGGGCCGGAAGTGTGAGACGCACAAGGATGACCAGCTGATC
 430 440 450 460 470 480
 +100 +110
 CysValAsnGluAsnGlyGlyCysGluGlnTyrCysSerAspHisThrGlyThrLysArg
 TGTGTGAACGAGAACGGCGGCTGTGAGCAGTACTGCAGTGACCACACGGGCACCAAGCGC
 490 500 510 520 530 540
 +120 +130
 SerCysArgCysHisGluGlyTyrSerLeuLeuAlaAspGlyValSerCysThrProThr
 TCCTGTGCGGTGCCACGAGGGGTACTCTGCTGTCGAGACGGGGTGTCTGTCACACCCACA
 550 560 570 580 590 600
 +140 +150
 ValGluTyrProCysGlyLysIleProIleLeuGluLysArgAsnAlaSerLysProGln
 GTTGAATATCCATGTGGAAAAATACCTATTCTAGAAAAAGAAATGCCAGCAAACCCCAA
 610 620 630 640 650 660

| | | | | | | | | | | | | | | | | | | | | | |
|--|------|------|------|------|------|--|--|--|--|------|--|--|--|--|--|--|--|--|--|--|------|
| | | | | | | | | | | +160 | | | | | | | | | | | +170 |
| GlyArgIleValGlyGlyLysValCysProLysGlyGluCysProTrpGlnValLeuLeu | | | | | | | | | | | | | | | | | | | | | |
| GGCCGAATTGTGGGGGGCAAGGTGTGCCCCAAAGGGGAGTGTCCATGGCAGGTCCTGTTG | | | | | | | | | | | | | | | | | | | | | |
| 670 | 680 | 690 | 700 | 710 | 720 | | | | | | | | | | | | | | | | |
| | | | | | | | | | | +180 | | | | | | | | | | | +190 |
| LeuValAsnGlyAlaGlnLeuCysGlyGlyThrLeuIleAsnThrIleTrpValValSer | | | | | | | | | | | | | | | | | | | | | |
| TTGGTGAATGGAGCTCAGTTGTGTGGGGGGACCCTGATCAACACCATCTGGGTGGTCTCC | | | | | | | | | | | | | | | | | | | | | |
| 730 | 740 | 750 | 760 | 770 | 780 | | | | | | | | | | | | | | | | |
| | | | | | | | | | | +200 | | | | | | | | | | | +210 |
| AlaAlaHisCysPheAspLysIleLysAsnTrpArgAsnLeuIleAlaValLeuGlyGlu | | | | | | | | | | | | | | | | | | | | | |
| GCGGCCCACTGTTTCGACAAAATCAAGAACTGGAGGAACCTGATCGCGGTGCTGGGCGAG | | | | | | | | | | | | | | | | | | | | | |
| 790 | 800 | 810 | 820 | 830 | 840 | | | | | | | | | | | | | | | | |
| | | | | | | | | | | +220 | | | | | | | | | | | +230 |
| HisAspLeuSerGluHisAspGlyAspGluGlnSerArgArgValAlaGlnValIleIle | | | | | | | | | | | | | | | | | | | | | |
| CACGACCTCAGCGAGCACGACGGGGATGAGCAGAGCCGGCGGGTGGCGCAGGTCATCATC | | | | | | | | | | | | | | | | | | | | | |
| 850 | 860 | 870 | 880 | 890 | 900 | | | | | | | | | | | | | | | | |
| | | | | | | | | | | +240 | | | | | | | | | | | +250 |
| ProSerThrTyrValProGlyThrThrAsnHisAspIleAlaLeuLeuArgLeuHisGln | | | | | | | | | | | | | | | | | | | | | |
| CCCAGCACGTACGTCCCGGGCACCAACCACGACATCGCGCTGCTCCGCCTGCACCAG | | | | | | | | | | | | | | | | | | | | | |
| 910 | 920 | 930 | 940 | 950 | 960 | | | | | | | | | | | | | | | | |
| | | | | | | | | | | +260 | | | | | | | | | | | +270 |
| ProValValLeuThrAspHisValValProLeuCysLeuProGluArgThrPheSerGlu | | | | | | | | | | | | | | | | | | | | | |
| CCCGTGGTCCTCACTGACCATGTGGTGCCCTCTGCCTGCCCGAACGGACGTTCTCTGAG | | | | | | | | | | | | | | | | | | | | | |
| 970 | 980 | 990 | 1000 | 1010 | 1020 | | | | | | | | | | | | | | | | |
| | | | | | | | | | | +280 | | | | | | | | | | | +290 |
| ArgThrLeuAlaPheValArgPheSerLeuValSerGlyTrpGlyGlnLeuLeuAspArg | | | | | | | | | | | | | | | | | | | | | |
| AGGACGCTGGCCTTCGTGCGCTTCTCATTGGTCAGCGGCTGGGGCCAGCTGCTGGACCGT | | | | | | | | | | | | | | | | | | | | | |
| 1030 | 1040 | 1050 | 1060 | 1070 | 1080 | | | | | | | | | | | | | | | | |
| | | | | | | | | | | +300 | | | | | | | | | | | +310 |
| GlyAlaThrAlaLeuGluLeuMetValLeuAsnValProArgLeuMetThrGlnAspCys | | | | | | | | | | | | | | | | | | | | | |
| GGCGCCACGGCCCTGGAGCTCATGGTCCTCAACGTGCCCCGGCTGATGACCCAGGACTGC | | | | | | | | | | | | | | | | | | | | | |
| 1090 | 1100 | 1110 | 1120 | 1130 | 1140 | | | | | | | | | | | | | | | | |
| | | | | | | | | | | +320 | | | | | | | | | | | +330 |
| LeuGlnGlnSerArgLysValGlyAspSerProAsnIleThrGluTyrMetPheCysAla | | | | | | | | | | | | | | | | | | | | | |
| CTGCAGCAGTCACGGAAGGTGGGAGACTCCCCAAATATCACGGAGTACATGTTCTGTGCC | | | | | | | | | | | | | | | | | | | | | |
| 1150 | 1160 | 1170 | 1180 | 1190 | 1200 | | | | | | | | | | | | | | | | |
| | | | | | | | | | | +340 | | | | | | | | | | | +350 |
| GlyTyrSerAspGlySerLysAspSerCysLysGlyAspSerGlyGlyProHisAlaThr | | | | | | | | | | | | | | | | | | | | | |
| GGCTACTCGGATGGCAGCAAGGACTCCTGCAAGGGGGACAGTGGAGGCCACATGCCACC | | | | | | | | | | | | | | | | | | | | | |
| 1210 | 1220 | 1230 | 1240 | 1250 | 1260 | | | | | | | | | | | | | | | | |
| | | | | | | | | | | +360 | | | | | | | | | | | +370 |
| HisTyrArgGlyThrTrpTyrLeuThrGlyIleValSerTrpGlyGlnGlyCysAlaThr | | | | | | | | | | | | | | | | | | | | | |
| CACTACCGGGGCACGTGGTACCTGACGGGCATCGTCAGCTGGGGCCAGGGCTGCGCAACC | | | | | | | | | | | | | | | | | | | | | |
| 1270 | 1280 | 1290 | 1300 | 1310 | 1320 | | | | | | | | | | | | | | | | |

| | | | | | | |
|--|------|------|------|------|------|------|
| | | | | +380 | | +390 |
| ValGlyHisPheGlyValTyrThrArgValSerGlnTyrIleGluTrpLeuGlnLysLeu | | | | | | |
| GTGGGCCACTTTGGGGTGTACACCAGGGTCTCCCAGTACATCGAGTGGCTGCAAAGCTC | | | | | | |
| 1330 | 1340 | 1350 | 1360 | 1370 | 1380 | |
| | | | +400 | | +406 | |
| MetArgSerGluProArgProGlyValLeuLeuArgAlaProPhePro*** | | | | | | |
| ATGCGCTCAGACCACGCCAGGAGTCCTCCTGCGAGCCCCATTTCCCTAGCCCAGCAGC | | | | | | |
| 1390 | 1400 | 1410 | 1420 | 1430 | 1440 | |
| CCTGGCCTGTGGAGAGAAAGCCAAGGCTGCGTCGAACTGTCTGGCACCAAATCCCATAT | | | | | | |
| 1450 | 1460 | 1470 | 1480 | 1490 | 1500 | |
| ATTCTTCTGCAGTTAATGGGGTAGAGGAGGGCATGGGAGGGAGGGAGAGGTGGGGAGGGA | | | | | | |
| 1510 | 1520 | 1530 | 1540 | 1550 | 1560 | |
| GACAGAGACAGAAACAGAGAGAGACAGAGACAGAGAGAGACTGAGGGAGAGACTCTGAGG | | | | | | |
| 1570 | 1580 | 1590 | 1600 | 1610 | 1620 | |
| ACATGGAGAGAGACTCAAAGAGACTCCAAGATTCAAAGAGACTAATAGAGACACAGAGAT | | | | | | |
| 1630 | 1640 | 1650 | 1660 | 1670 | 1680 | |
| GGAATAGAAAAGATGAGAGGCAGAGGCAGACAGGCGCTGGACAGAGGGGCAGGGGAGTGC | | | | | | |
| 1690 | 1700 | 1710 | 1720 | 1730 | 1740 | |
| CAAGGTTGTCCTGGAGGCAGACAGCCCAGCTGAGCCTCCTTACCTCCCTTCAGCCAAGCC | | | | | | |
| 1750 | 1760 | 1770 | 1780 | 1790 | 1800 | |
| CCACCTGCACGTGATCTGCTGGCCCTCAGGCTGCTGCTCTGCCTTCATTGCTGGAGACAG | | | | | | |
| 1810 | 1820 | 1830 | 1840 | 1850 | 1860 | |
| TAGAGGCATGAACACACATGGATGCACACACACACACGCCAATGCACACACACAGAGATA | | | | | | |
| 1870 | 1880 | 1890 | 1900 | 1910 | 1920 | |
| TGCACACACACGGATGCACACACAGATGGTCACACAGAGATACGCAAACACACCGATGCA | | | | | | |
| 1930 | 1940 | 1950 | 1960 | 1970 | 1980 | |
| CACGCACATAGAGATATGCACACACAGATGCACACACAGATATACACATGGATGCACGCA | | | | | | |
| 1990 | 2000 | 2010 | 2020 | 2030 | 2040 | |
| CATGCCAATGCACGCACACATCAGTGCACACGGATGCACAGAGATATGCACACACCGATG | | | | | | |
| 2050 | 2060 | 2070 | 2080 | 2090 | 2100 | |
| TGCGCACACACAGATATGCACACACATGGATGAGCACACACACACCAAGTGCGCACACAC | | | | | | |
| 2110 | 2120 | 2130 | 2140 | 2150 | 2160 | |
| ACCGATGTACACACACAGATGCACACACAGATGCACACACACCGATGCTGACTCCATGTG | | | | | | |
| 2170 | 2180 | 2190 | 2200 | 2210 | 2220 | |
| TGCTGTCTCTGAAGGCGGTTGTTTAGCTCTCACTTTTCTGGTTCTTATCCATTATCATC | | | | | | |
| 2230 | 2240 | 2250 | 2260 | 2270 | 2280 | |
| TTCACTTCAGACAATTGAGAAGCATCACCATGCATGGTGGCGAATGCCCCAAACTCTCC | | | | | | |
| 2290 | 2300 | 2310 | 2320 | 2330 | 2340 | |
| CCCAAATGTATTTCTCCCTTCGCTGGGTGCCGGGCTGCACAGACTATTCCCACCTGCTT | | | | | | |
| 2350 | 2360 | 2370 | 2380 | 2390 | 2400 | |

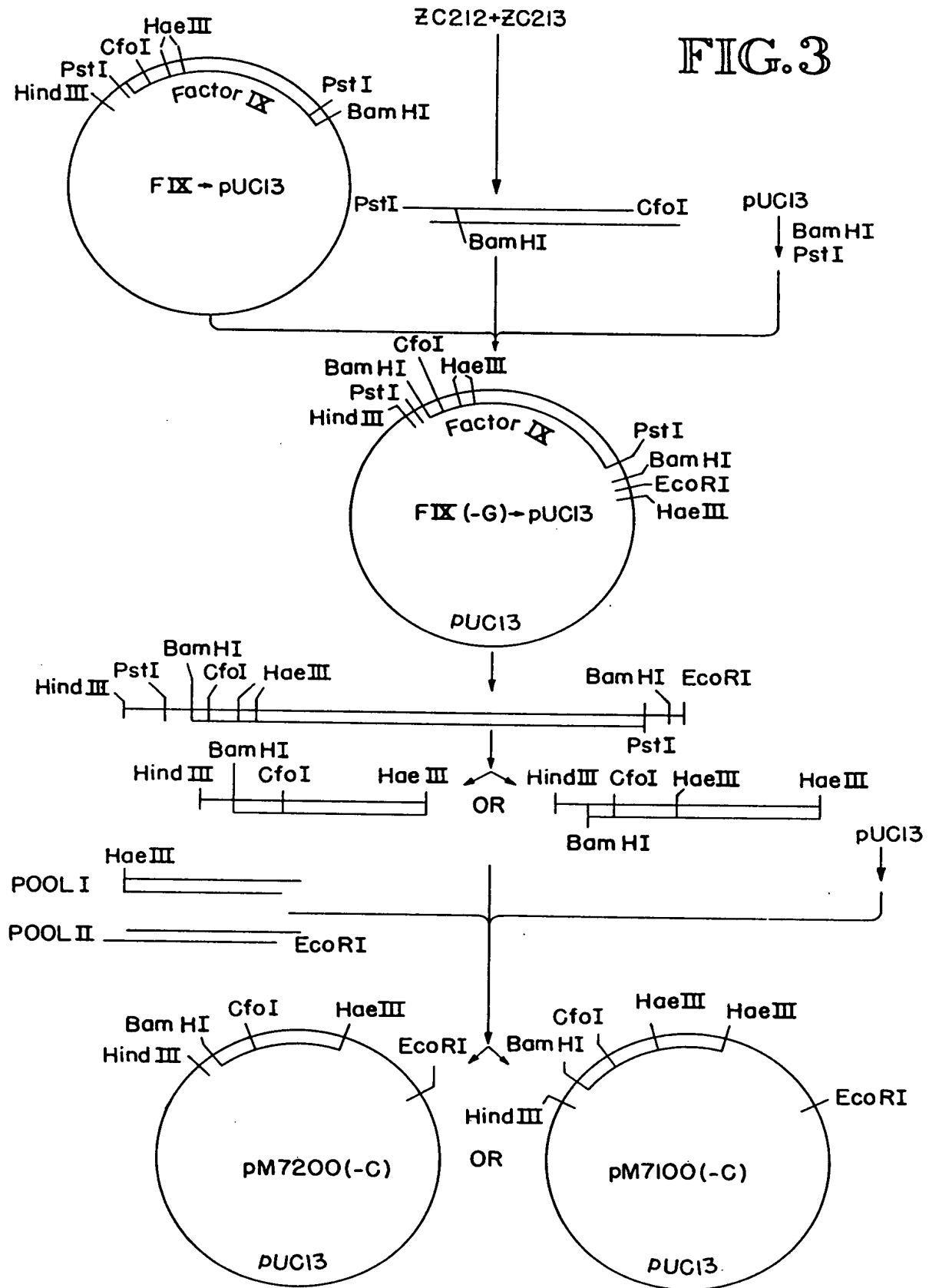
CCCAGCTTCACAATAAACGGCTGCGTCTCCTCCGCACACCTGTGGTGCCTGCCACCCAAA
2410 2420 2430 2440 2450 2460

AAAAAAAAAAAAAAAAAAAA
2470 2480

FIG. 2B

| | | | | | |
|---------------------|---------|-----------|---------|---------|-----------------|
| From cDNA | 1 | 10 | 20 | 30 | 36 |
| | XXXXXX | XXXXXX | XXXXXX | XXXXXX | XXXXRT |
| Amino Acid Sequence | AAFLY | YLRPGSL | YRYCKYY | QCIFY | ARYIFXXXXX |
| | | | | | |
| cDNA | 40 | 50 | 60 | 70 | |
| | KLFWISY | SDGDQCA | SSPCQNG | SGCKDQL | QSYICFCL |
| Amino Acid Sequence | LFWISY | SDGDQCA | SSPCQNG | SGCKDQL | Q ICFCL |
| | | | | | |
| cDNA | 80 | 90 | 100 | | |
| | PAFEGRN | CECHKDDQL | ICVNE | NGGCEQY | CSDHTGTK |
| Amino Acid Sequence | PAFEGRN | CECHKDDQL | | | CSDHTGT |
| | | | | | |
| cDNA | 110 | 120 | 130 | 140 | |
| | RSCRCH | EGYSL | LADGV | SCTPT | VEYPCGKIPILEKRN |
| Amino Acid Sequence | RSCRCH | EGYSL | LADGV | SCTPT | VEY EKR() |
| | | | | | |
| cDNA | 150 | | | | |
| | ASKPQGR | | | | |
| Amino Acid Sequence | ASKPQGR | | | | |
| | | | | | |

FIG. 3



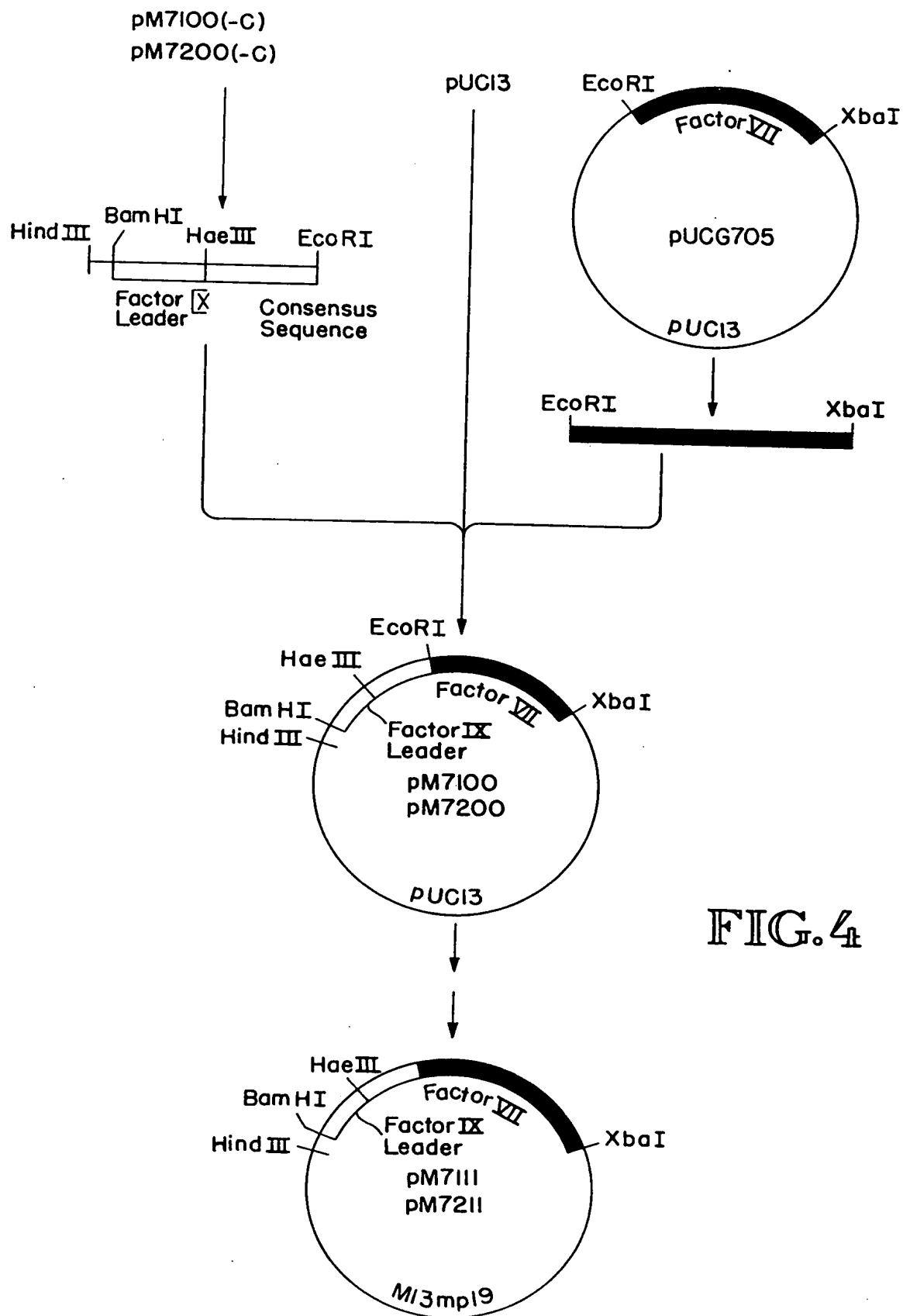


FIG. 4

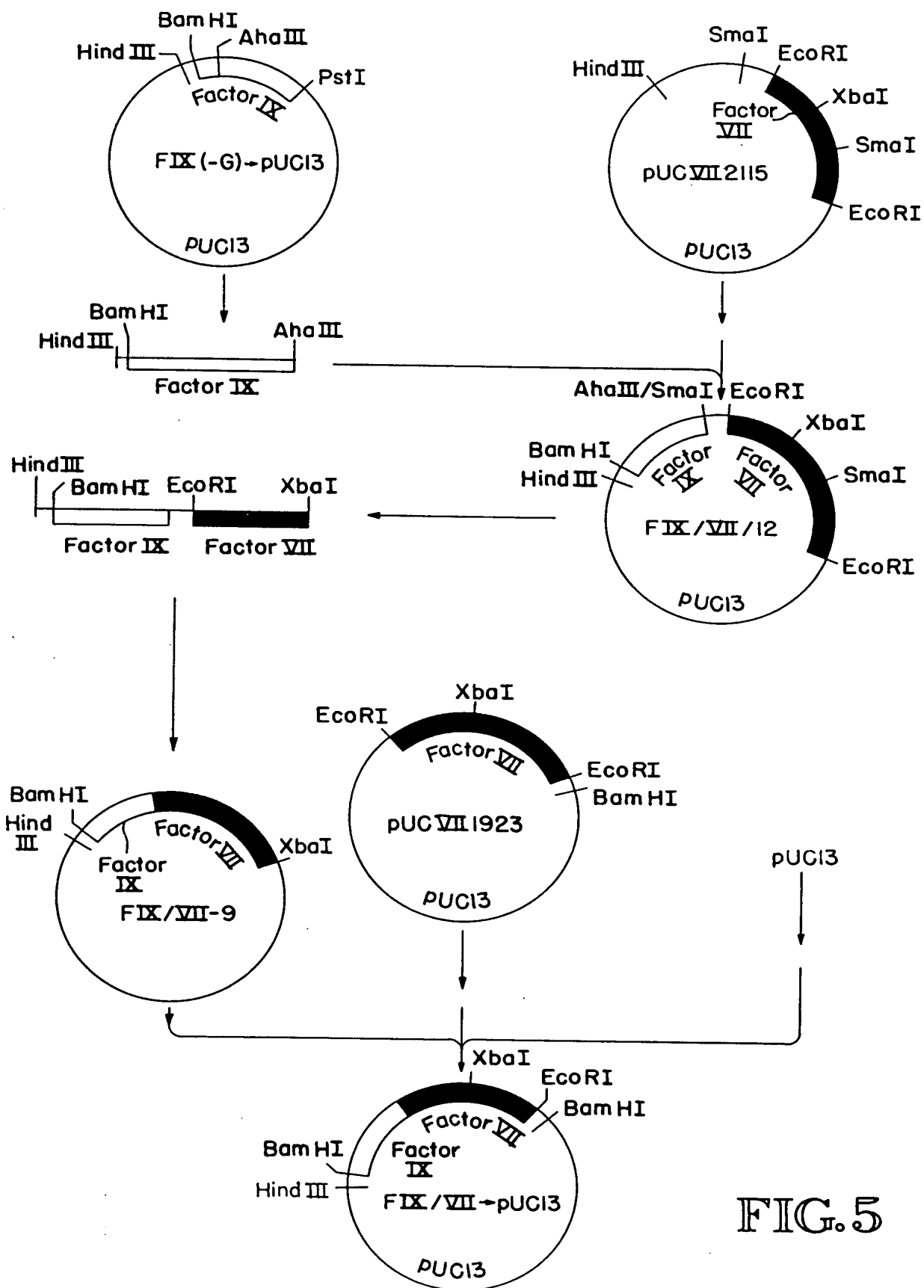


FIG.5

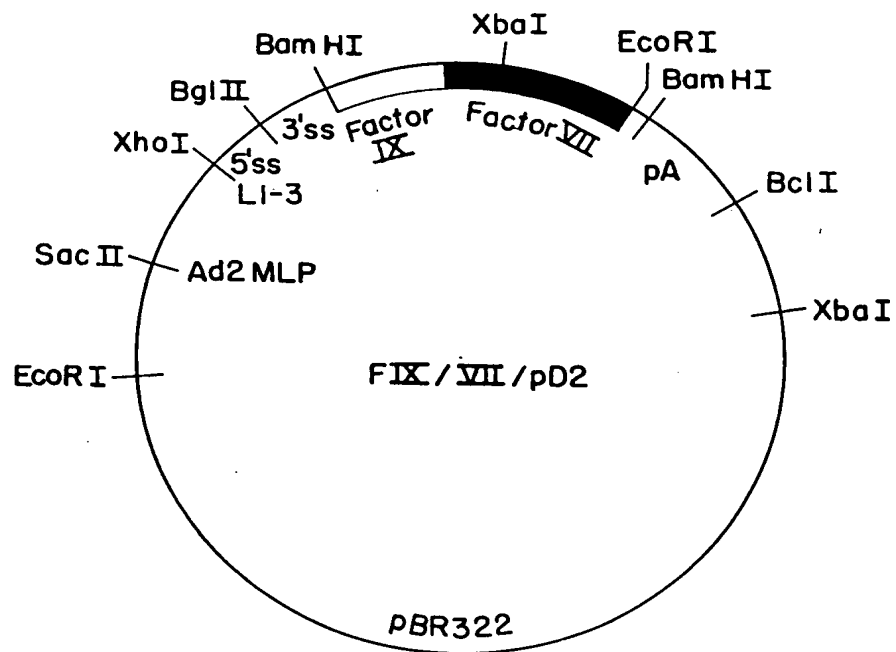


FIG. 6

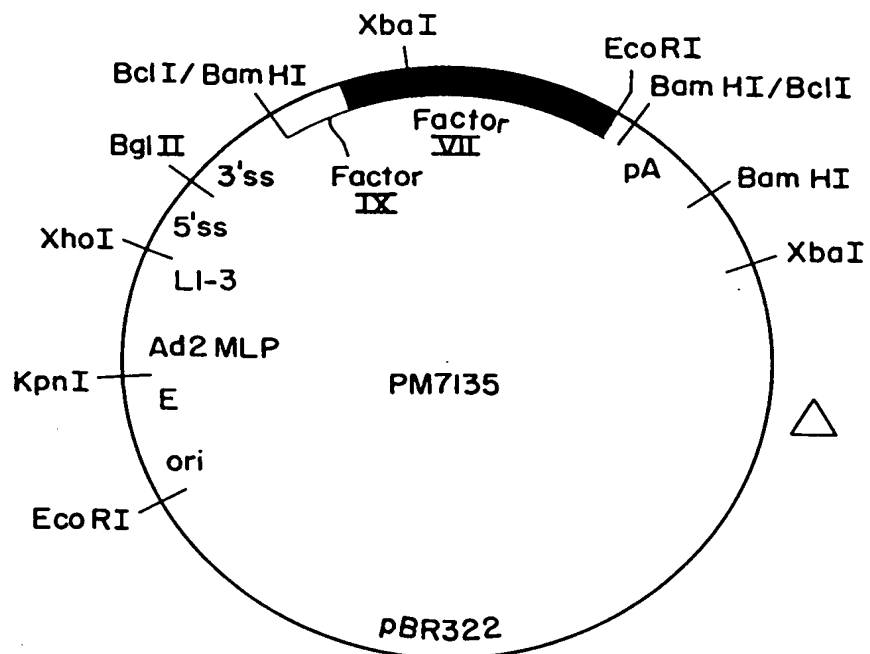


FIG. 8

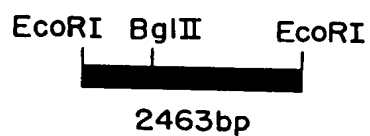
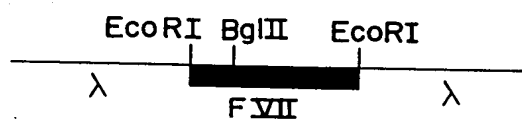
FIG. 7

| | | | | | | | | | | | | | | | | | |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|
| | | | | | | 21 | | | | | | | 36 | | | | |
| GGATCC | ATG | CAG | CGC | GTG | AAC | ATG | ATC | ATG | GCA | GAA | TCA | CCA | GGC | | | | |
| | MET | Gln | Arg | Val | Asn | MET | Ile | MET | Ala | Glu | Ser | Pro | Gly | | | | |
| | | | | | | 66 | | | | | | | 81 | | | | |
| CTC | ATC | ACC | ATC | TGC | CTT | TTA | GGA | TAT | CTA | CTC | AGT | GCT | GAA | TGT | | | |
| Leu | Ile | Thr | Ile | Cys | Leu | Leu | Gly | Tyr | Leu | Leu | Ser | Ala | Glu | Cys | | | |
| | | | | | | 96 | | | | | | | 111 | | | | |
| ACA | GTT | TTT | CTT | GAT | CAT | GAA | AAC | GCC | AAC | AAA | ATT | CTG | AAT | CGG | | | |
| Thr | Val | Phe | Leu | Asp | His | Glu | Asn | Ala | Asn | Lys | Ile | Leu | Asn | Arg | | | |
| | | | | | | 141 | | | | | | | 156 | | | | |
| CCA | AAG | AGG | TAT | AAT | TCA | GGT | AAA | TTG | GAA | GAG | TTT | GTT | CAA | GGG | | | |
| Pro | Lys | Arg | Tyr | Asn | Ser | Gly | Lys | Leu | Glu | Glu | Phe | Val | Gln | Gly | | | |
| | | | | | | 186 | | | | | | | 201 | | | | |
| AAC | CTT | GAG | AGA | GAA | TGT | ATG | GAA | GAA | AAG | TGT | AGT | TTT | GAA | GAA | | | |
| Asn | Leu | Glu | Arg | Glu | Cys | MET | Glu | Glu | Lys | Cys | Ser | Phe | Glu | Glu | | | |
| | | | | | | 231 | | | | | | | 246 | | | | |
| GCA | CGA | GAA | GTT | TTT | GAA | AAC | ACT | GAA | AGA | ACA | AAG | CTG | TTC | TGG | | | |
| Ala | Arg | Glu | Val | Phe | Glu | Asn | Thr | Glu | Arg | Thr | Lys | Leu | Phe | Trp | | | |
| | | | | | | 276 | | | | | | | 291 | | | | |
| ATT | TCT | TAC | AGT | GAT | GGG | GAC | CAG | TGT | GCC | TCA | AGT | CCA | TGC | CAG | | | |
| Ile | Ser | Tyr | Ser | Asp | Gly | Asp | Gln | Cys | Ala | Ser | Ser | Pro | Cys | Gln | | | |
| | | | | | | 321 | | | | | | | 336 | | | | |
| AAT | GGG | GGC | TCC | TGC | AAG | GAC | CAG | CTC | CAG | TCC | TAT | ATC | TGC | TTC | | | |
| Asn | Gly | Gly | Ser | Cys | Lys | Asp | Gln | Leu | Gln | Ser | Tyr | Ile | Cys | Phe | | | |
| | | | | | | 366 | | | | | | | 381 | | | | |
| TGC | CTC | CCT | GCC | TTC | GAG | GGC | CGG | AAC | TGT | GAG | ACG | CAC | AAG | GAT | | | |
| Cys | Leu | Pro | Ala | Phe | Glu | Gly | Arg | Asn | Cys | Glu | Thr | His | Lys | Asp | | | |
| | | | | | | 411 | | | | | | | 426 | | | | |
| GAC | CAG | CTG | ATC | TGT | GTG | AAC | GAG | AAC | GGC | GGC | TGT | GAG | CAG | TAC | | | |
| Asp | Glu | Leu | Ile | Cys | Val | Asn | Glu | Asn | Gly | Gly | Cys | Glu | Gln | Tyr | | | |
| | | | | | | 456 | | | | | | | 471 | | | | |
| TGC | AGT | GAC | CAC | ACG | GGC | ACC | AAG | CGC | TCC | TGT | CGG | TGC | CAC | GAG | | | |
| Cys | Ser | Asp | His | Thr | Gly | Thr | Lys | Arg | Ser | Cys | Arg | Cys | His | Glu | | | |
| | | | | | | 501 | | | | | | | 516 | | | | |
| GGG | TAC | TCT | CTG | CTG | GCA | GAC | GGG | GTG | TCC | TGC | ACA | CCC | ACA | GTT | | | |
| Gly | Tyr | Ser | Leu | Leu | Ala | Asp | Gly | Val | Ser | Cys | Thr | Pro | Thr | Val | | | |
| | | | | | | 546 | | | | | | | 561 | | | | |
| GAA | TAT | CCA | TCT | GGA | AAA | ATA | CCT | ATT | CTA | GAA | AAA | AGA | AAT | GCC | | | |
| Glu | Tyr | Pro | Cys | Gly | Lys | Ile | Pro | Ile | Leu | Glu | Lys | Arg | Asn | Ala | | | |
| | | | | | | 591 | | | | | | | 606 | | | | |
| AGC | AAA | CCC | CAA | GGC | CGA | ATT | GTG | GGG | GGC | AAG | GTG | TGC | CCC | AAA | | | |
| Ser | Lys | Pro | Gln | Gly | Arg | Ile | Val | Gly | Gly | Lys | Val | Cys | Pro | Lys | | | |

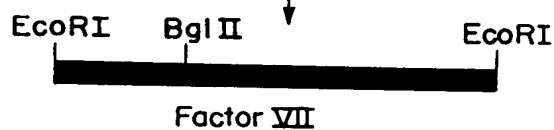
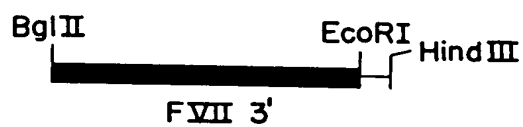
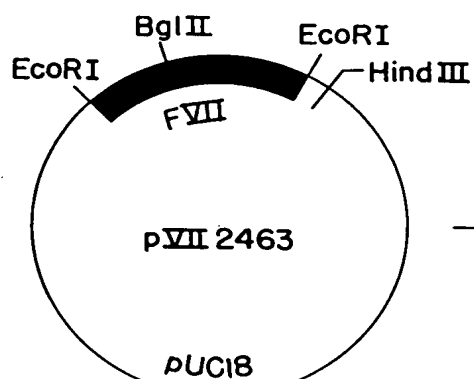
| | | | | | |
|---|---|------|---|------|---|
| 636 | GGG GAG TGT CCA TGG CAG GTC CTG TTG TTG GTG AAT GGA GCT CAG | 651 | CTG TTG TTG GTG AAT GGA GCT CAG | 666 | AAT GGA GCT CAG |
| Gly Glu Cys Pro Trp Gln Val Leu Leu Leu Val Asn Gly Ala Gln | | | | | |
| 681 | TTG TGT GGG GGG ACC CTG ATC AAC ACC ATC TGG GTG GTC TCC GCG | 696 | AAC ACC ATC TGG GTG GTC TCC GCG | 711 | GTG GTC TCC GCG |
| Leu Cys Gly Gly Thr Leu Ile Asn Thr Ile Trp Val Val Ser Ala | | | | | |
| 726 | GCC CAC TGT TTC GAC AAA ATC AAG AAC TGG AGG AAC CTG ATC GCG | 741 | AAG AAC TGG AGG AAC CTG ATC GCG | 756 | AAC CTG ATC GCG |
| Ala His Cys Phe Asp Lys Ile Lys Asn Trp Arg Asn Leu Ile Ala | | | | | |
| 771 | GTG CTG GGC GAG CAC GAC CTC AGC GAG CAC GAC GGG GAT GAG CAG | 786 | AGC GAG CAC GAC GGG GAT GAG CAG | 801 | GAT GAG CAG |
| Val Leu Gly Glu His Asp Leu Ser Glu His Asp Gly Asp Glu Gln | | | | | |
| 816 | AGC CGG CGG GTG GCG CAG GTC ATC ATC CCC AGC ACG TAC GTC CCG | 831 | GTC ATC ATC CCC AGC ACG TAC GTC CCG | 846 | ACG TAC GTC CCG |
| Ser Arg Arg Val Ala Gln Val Ile Ile Pro Ser Thr Tyr Val Pro | | | | | |
| 861 | GGC ACC ACC AAC CAC GAC ATC GCG CTG CTC CGC CTG CAC CAG CCC | 876 | GCG CTG CTC CGC CTG CAC CAG CCC | 891 | CTG CAC CAG CCC |
| Gly Thr Thr Asn His Asp Ile Ala Leu Leu Arg Leu His Gln Pro | | | | | |
| 906 | GTG GTC CTC ACT GAC CAT GTG GTG CCC CTC TGC CTG CCC GAA CGG | 921 | GTG GTG CCC CTC TGC CTG CCC GAA CGG | 936 | CTG CCC GAA CGG |
| Val Val Leu Thr Asp His Val Val Pro Leu Cys Leu Pro Glu Arg | | | | | |
| 951 | ACG TTC TCT GAG AGG ACG CTG GCC TTC GTG CGC TTC TCA TTG GTC | 966 | CTG GCC TTC GTG CGC TTC TCA TTG GTC | 981 | TTC TCA TTG GTC |
| Thr Phe Ser Glu Arg Thr Leu Ala Phe Val Arg Phe Ser Leu Val | | | | | |
| 996 | AGC GGC TGG GGC CAG CTG CTG GAC CGT GGC GCC ACG GCC CTG GAG | 1011 | CTG GAC CGT GGC GCC ACG GCC CTG GAG | 1026 | ACG GCC CTG GAG |
| Ser Gly Trp Gly Gln Leu Leu Asp Arg Gly Ala Thr Ala Leu Glu | | | | | |
| 1041 | CTC ATG GTC CTC AAC GTG CCC CGG CTG ATG ACC CAG GAC TGC CTG | 1056 | CTG CCC CGG CTG ATG ACC CAG GAC TGC CTG | 1071 | ACC CAG GAC TGC CTG |
| Leu MET Val Leu Asn Val Pro Arg Leu MET Thr Gln Asp Cys Leu | | | | | |
| 1086 | CAG CAG TCA CGG AAG GTG GGA GAC TCC CCA AAT ATC ACG GAG TAC | 1101 | GAG GGA GAC TCC CCA AAT ATC ACG GAG TAC | 1116 | ATC ACG GAG TAC |
| Gln Gln Ser Arg Lys Val Gly Asp Ser Pro Asn Ile Thr Glu Tyr | | | | | |
| 1131 | ATG TTC TGT GCC GGC TAC TCG GAT GGC AGC AAG GAC TCC TGC AAG | 1146 | TAC TCG GAT GGC AGC AAG GAC TCC TGC AAG | 1161 | GAC TCC TGC AAG |
| MET Phe Cys Ala Gly Tyr Ser Asp Gly Ser Lys Asp Ser Cys Lys | | | | | |
| 1176 | GGG GAC AGT GGA GGC CCA CAT GCC ACC CAC TAC CGG GGC ACG TGG | 1191 | GAC AGT GGA GGC CCA CAT GCC ACC CAC TAC CGG GGC ACG TGG | 1206 | AGT GGA GGC CCA CAT GCC ACC CAC TAC CGG GGC ACG TGG |
| Gly Asp Ser Gly Gly Pro His Ala Thr His Tyr Arg Gly Thr Trp | | | | | |
| 1221 | TAC CTG ACG GGC ATC GTC AGC TGG GGC CAG GGC TGC GCA ACC GTG | 1236 | CTG ACG GGC ATC GTC AGC TGG GGC CAG GGC TGC GCA ACC GTG | 1251 | ACG GGC ATC GTC AGC TGG GGC CAG GGC TGC GCA ACC GTG |
| Tyr Leu Thr gly Ile Val Ser Trp Gly Gln Gly Cys Ala Thr Val | | | | | |

| | | |
|---|------|------|
| 1266 | 1281 | 1296 |
| GGC CAC TTT GGG GTG TAC ACC AGG GTC TCC CAG TAC ATC GAG TGG | | |
| Gly His Phe Gly Val Tyr Thr Arg Val Ser Gln Tyr Ile Glu Trp | | |
| 1311 | 1326 | 1341 |
| CTG CAA AAG CTC ATG CGC TCA GAG CCA CGC CCA GGA GTC CTC CTG | | |
| Leu Gln Lys Leu MET Arg Ser Glu Pro Arg Pro Gly Val Leu Leu | | |
| 1356 | 1378 | 1388 |
| CGA GCC CCA TTT CCC TAG CCCAGCAGCC CTGGCCTGTG GAGAGAAAGC | | |
| Arg Ala Pro Phe Pro | | |
| 1408 | 1418 | 1428 |
| CAAGGCTGCG TCGAACTGTC CTGGCACCAA ATCCCATATA TTCTTCTGCA | | |
| 1458 | 1468 | 1478 |
| GTTAATGGGG TAGAGGAGGG CATGGGAGGG AGGGAGAGGT GGGGAGGGAG | | |
| 1508 | 1518 | 1528 |
| ACAGAGACAG AACACAGAGAG AGACAGAGAC AGAGAGAGAC TGAGGGGAGAG | | |
| 1558 | 1568 | 1578 |
| ACTCTGAGGA CCATGGAGAG AGACTCAAAG AGACTCCAAG ATTCAAAGAG | | |
| 1608 | 1618 | 1628 |
| ACTAATAGAG ACACAGAGAT GGAATAGAAA AGATGAGAGG CAGAGGCAGA | | |
| 1658 | 1668 | 1678 |
| CAGGCGCTGG ACAGAGGGGC AGGGGAGTGC CAAGGTTGTC CTGGAGGCAG | | |
| 1708 | 1718 | 1728 |
| ACAGCCCAGC TGAGCCTCCT TACCTCCCTT CAGCCAAGCC CCACCTGCAC | | |
| 1758 | 1768 | 1778 |
| GTGATCTGCT GGCCCTCAGG CTGCTGCTCT GCCTTCATTG CTGGAGACAG | | |
| 1808 | 1818 | 1828 |
| TAGAGGCATG ACACACATGG ATGCACACAC ACACACGCCA TGCACACACA | | |
| 1858 | 1868 | 1878 |
| CAGAGATATG CACACACACG GATGCACACA CAGATGGTCA CACAGAGTAC | | |
| 1908 | 1918 | 1928 |
| GCAAACACAC CGATGCACAC GCACATAGAG ATATGCACAC ACAGATGCAC | | |

| | | | | |
|------------|------------|------------|------------|------------|
| 1958 | 1968 | 1978 | 1988 | 1998 |
| ACACAGATAT | ACACATGGAG | TGCACGCACA | TGCCAATGCA | CGCACACATC |
| 2008 | 2018 | 2028 | 2038 | 2048 |
| AGTGACACAG | GATGCACAGA | GATATGCACA | CACCGATGTG | CGCACACACA |
| 2058 | 2068 | 2078 | 2088 | 2098 |
| GATATGCACA | CACATGGATG | AGCACACACA | CACCAAGTGC | GCACACACAC |
| 2108 | 2118 | 2128 | 2138 | 2148 |
| CGATGTACAC | ACAGATGCAC | ACACAGATGC | ACACACACCG | ATGCTGACTC |
| 2158 | 2168 | 2178 | 2188 | 2198 |
| CATGTGTGCT | GTCCTCTGAA | GGCGGTTGTT | TAGCTCTCAC | TTTTCTGGTT |
| 2208 | 2218 | 2228 | 2238 | 2248 |
| CTTATCCATT | ATCATCTTCA | CTTCAGACAA | TTCAGAAGCA | TCACCATGCA |
| 2258 | 2268 | 2278 | 2288 | 2298 |
| TGGTGGCGAA | TGCCCCCAA | CTCTCCCCCA | AATGTATTTC | TCCCTTCGCT |
| 2308 | 2318 | 2328 | 2338 | 2348 |
| GGGTGCCGGG | CTGCACAGAC | TATTCCCCAC | CTGCTTCCCA | GCTTCACAAT |
| 2358 | 2368 | 2378 | 2388 | 2398 |
| AAACGGCTGC | GTCTCCTCGC | AAAAAAAAAA | AAAAAAAAAA | AAAAAAAAAA |
| 2408 | 2418 | 2428 | 2438 | |
| AAAAAAAAAA | AAGGAATTCT | AGCTCGGTAC | CCGGGGATCC | |



+pUC18



TO MAMMALIAN CELL
EXPRESSION VECTOR

FIG. 9

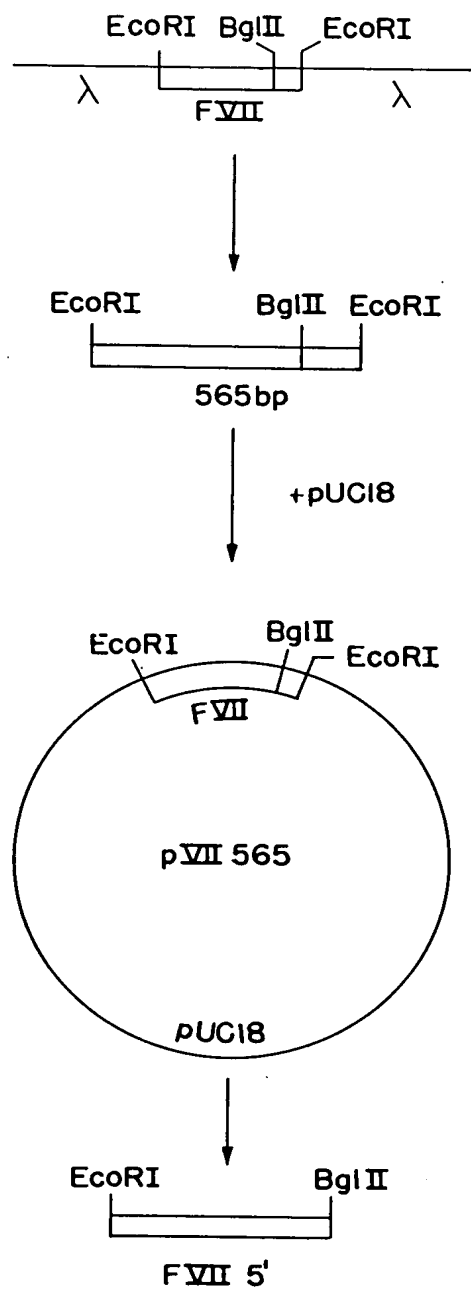


FIG. 10

FIG. 11

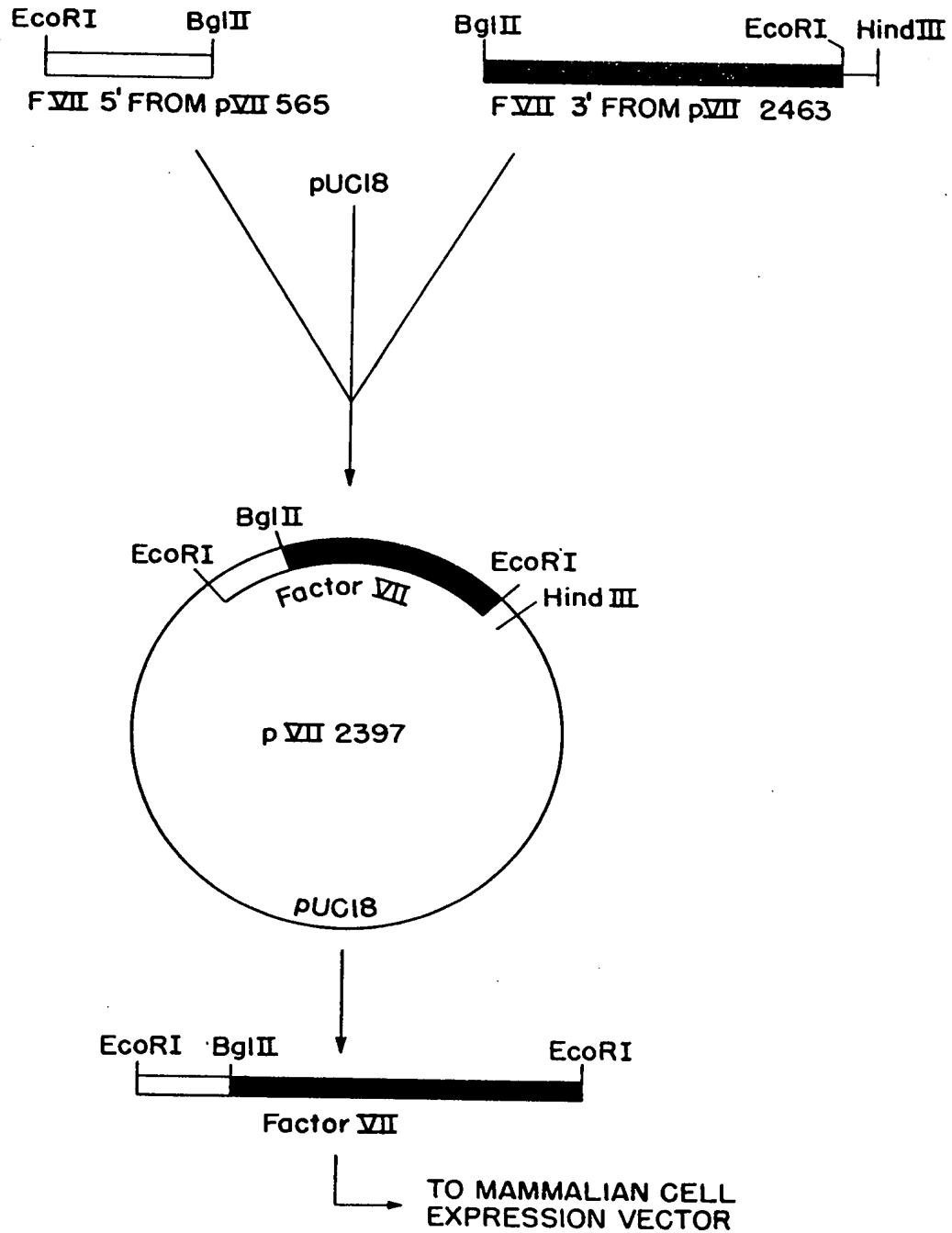


FIG. 12

